

WHAT IS CLAIMED IS:

1. A computer storage system, comprising:

a plurality of disk drives for storing distributed parity groups, each distributed parity group comprising storage blocks, said storage blocks comprising one or more data blocks and a parity block, each of said storage blocks stored on a separate disk drive such that no two storage blocks from a given parity group reside on the same disk drive;

file system metadata to describe a location of each of said storage blocks by specifying a disk identifier and a logical block identifier of a first logical block of each storage block; and

an allocation module configured to allocate space for a new distributed parity group on two or more disk drives in said plurality of disk drives, said allocation module allocating space for each storage block of said new distributed parity group from free space on each of said two or more disk drives.

2. The computer storage system of Claim 1, said allocation module further configured to recognize a new disk drive hot-swapped into said plurality of disk drives during file system operation and to use said new disk drive to store one or more storage blocks.

3. The computer storage system of Claim 1, said allocation module further configured to recognize a new disk drive hot-swapped into said plurality of disk drives during file system operation and to use any free logical blocks on said new disk drive to store one or more storage blocks.

4. The computer storage system of Claim 3, wherein said new disk drive is larger in capacity than any drive in said plurality of disk drives.

5. The computer storage system of Claim 1, wherein a size of a first distributed parity group is larger than a size of a second distributed parity group within a first file.

6. The computer storage system of Claim 1, wherein said new disk drive is provided to a Fibre Channel network.

7. The computer storage system of Claim 1, wherein a file is organized as one or more distributed parity groups of varying size.

8. The computer storage system of Claim 1, wherein an extent of a first distributed parity group of a file is larger than an extent of a second distributed parity group of said file.

9. The computer storage system of Claim 1, further comprising a load-balancing module to move one or more distributed parity groups from disk to disk to improve load balancing of said plurality of disks.

10. The computer storage system of Claim 1, further comprising a load-balancing module to move one or more distributed parity groups from disk to disk to improve capacity balancing of said plurality of disks.

11. The computer storage system of Claim 1, further comprising a load-balancing module to move one or more distributed parity groups from disk to disk to improve load balancing of said plurality of disks when a new disk is added to said plurality of disks.

12. The computer storage system of Claim 1, further comprising a load-balancing module to move one or more distributed parity groups from disk to disk to improve capacity balancing of said plurality of disks when a new disk is added to said plurality of disks.

13. A method for managing a storage array, comprising:

recognizing a new storage device;

adding said new storage device to a list of previously-available storage devices to produce a list of currently-available storage devices;

moving a selected storage block from a selected parity group to said new storage device;

updating file system metadata to reflect the new location of said selected storage device such that clients accessing a file containing said selected storage block can continue to access data in said selected storage block.

14. The method of Claim 13, wherein said updating comprises specifying a disk identifier and a logical block identifier.

15. The method of Claim 13, wherein said new storage device comprises a new disk drive.

16. The method of Claim 13, wherein said new disk drive is larger in capacity than other disk drives in said list of previously-available storage devices, and wherein said selected storage block can be stored on any unused logical block of said new disk drive.

5

10
15
20
25
30
35
40
45
50
55
60
65
70
75
80
85
90
95
100
105
110
115
120
125
130
135
140
145
150
155
160
165
170
175
180
185
190
195
200
205
210
215
220
225
230
235
240
245
250
255
260
265
270
275
280
285
290
295
300
305
310
315
320
325
330
335
340
345
350
355
360
365
370
375
380
385
390
395
400
405
410
415
420
425
430
435
440
445
450
455
460
465
470
475
480
485
490
495
500
505
510
515
520
525
530
535
540
545
550
555
560
565
570
575
580
585
590
595
600
605
610
615
620
625
630
635
640
645
650
655
660
665
670
675
680
685
690
695
700
705
710
715
720
725
730
735
740
745
750
755
760
765
770
775
780
785
790
795
800
805
810
815
820
825
830
835
840
845
850
855
860
865
870
875
880
885
890
895
900
905
910
915
920
925
930
935
940
945
950
955
960
965
970
975
980
985
990
995